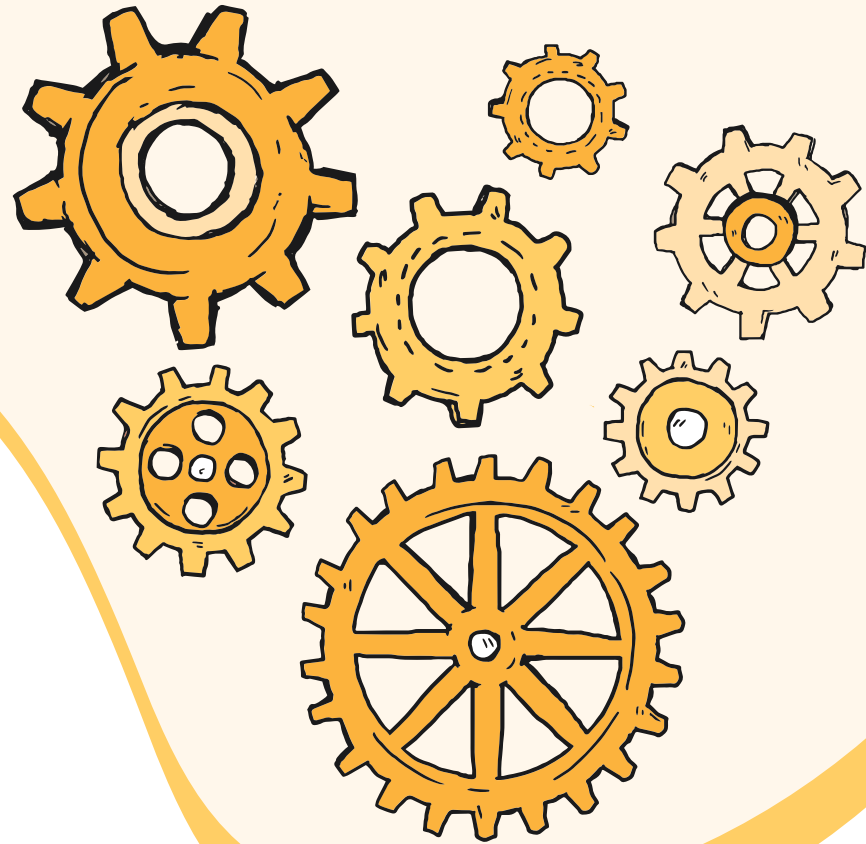
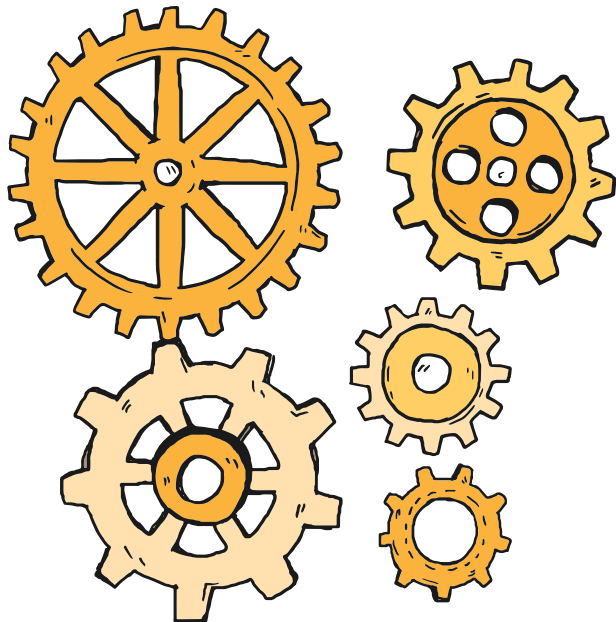


DBM1 - Databases Presentation

By Luca Bova, Chad Long &
Jordan Ukawoko





kaggle

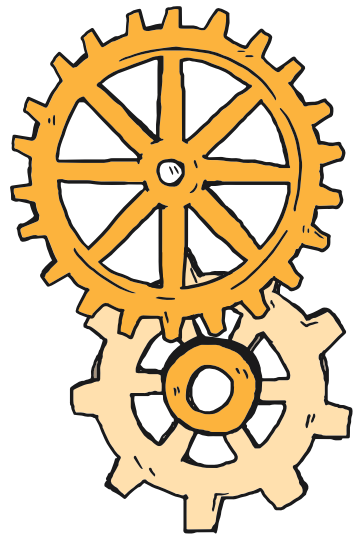
Video Game Market Dataset

- Found On Kaggle
- This dataset is a web scrape from VGChartz with thousands of entries ranging to details about sales of each games, their genre and the name of the publisher
- We modified the dataset to include games from different platforms
- The dataset gives us a legitimate and strong sample size

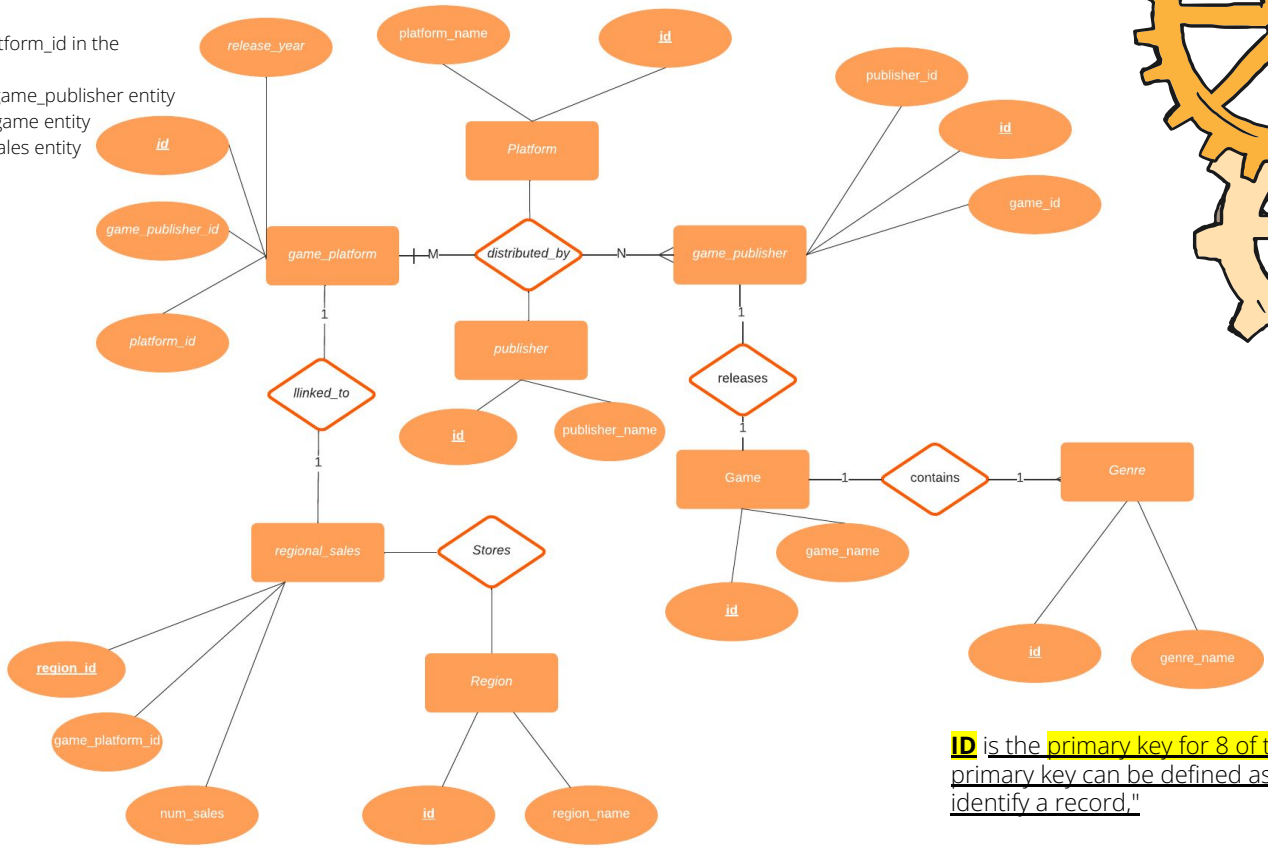
Foreign keys also exist within this relationship schema.

Example of the foreign keys are below. A foreign Key is a field in one table that refers to the primary key of another table

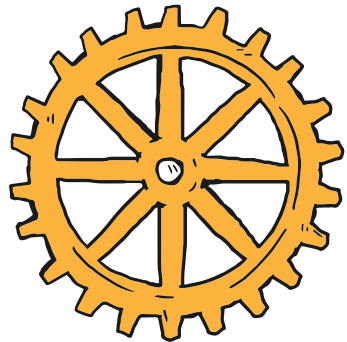
ER Diagram/Model



- game_publisher_id and game_platform_id in the game_platform entity
- game_id and publisher_id in the game_publisher entity
- genre_id and game_name in the game entity
- game_platform_id in the region_sales entity



For example - **game_platform_id** is a foreign key in the region_sales entity because **id** is the primary key in the game_platform entity.



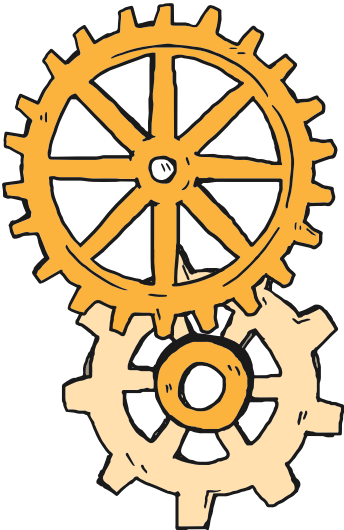
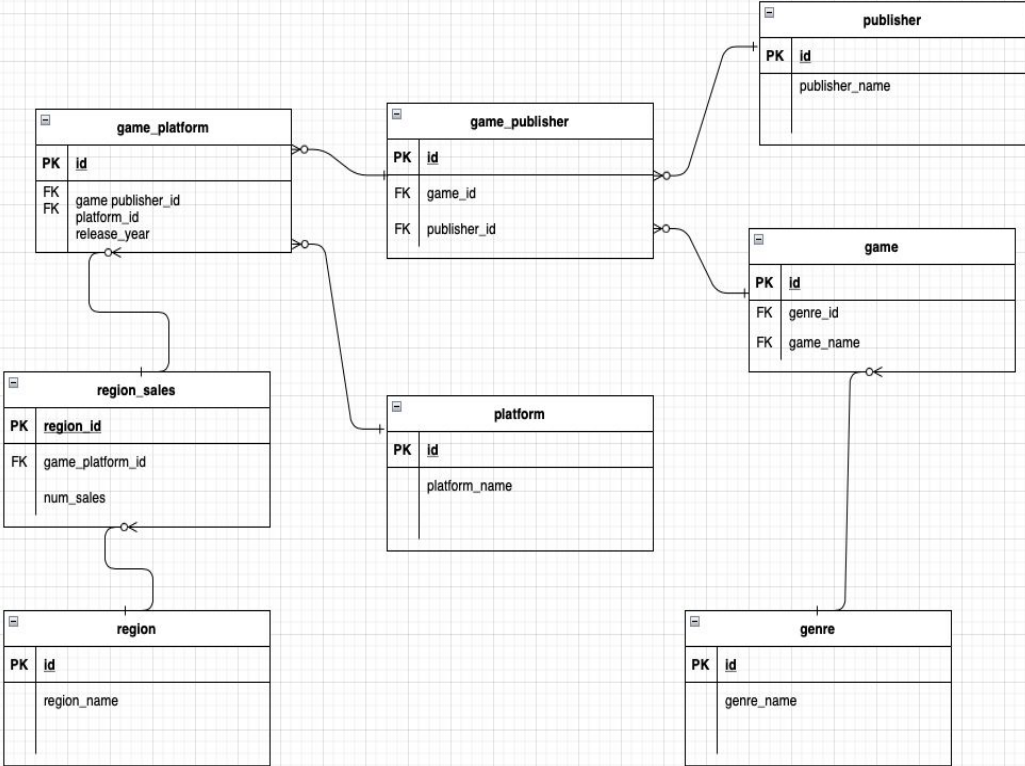
ID is the primary key for 8 of these entities. A primary key can be defined as "which attributes identify a record."

Foreign keys also exist within this relationship schema. Example of the foreign keys are below.

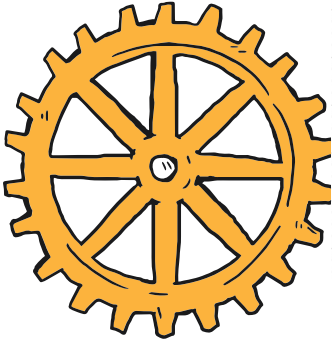
List of Entities

- game_publisher_id and game_platform_id in the game_platform entity
- game_id and publisher_id in the game_publisher entity
- genre_id and game_name in the game entity
- game_platform_id in the region_sales entity

For example - **game_platform_id** is a foreign key in the **region_sales** entity because **id** is the primary key in the game_platform entity.



Relationships **include 1:1 and M:N relationships.** More elaboration in report/live presentation





Natural Language Query



How many games were released in 2016?



Group the games from highest selling



What region accounts for the majority of sales?



List the games released by Activision



List the games for the Nintendo Wii Platform



Find the average number of sales of games released in North

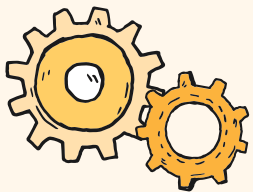


America

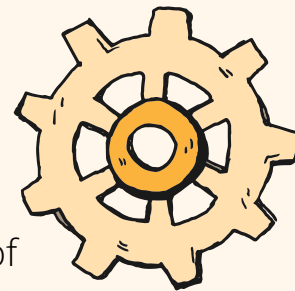


List the games that sold more than 10 million copies





Answers to Queries



How many **games** released in **2016**?

```
SELECT COUNT (release_year)
FROM game_pl
WHERE release_year = 2016
```

Group the games from **highest** selling

```
SELECT
g.game_name,
pl.platform_name,
gp.release_year,
pub.publisher_name,
SUM(rs.num_sales) AS global_sales
FROM regional_sales rs
INNER JOIN region r ON rs.region_id = r.id
INNER JOIN game_pl gp ON rs.game_platform_id
= gp.id
INNER JOIN game_publisher gpub ON
gp.game_publisher_id = gpub.id
INNER JOIN games g ON gpub.game_id = g.id
INNER JOIN platform pl ON gp.platform_id = pl.id
INNER JOIN publisher pub ON gpub.publisher_id
= pub.id
GROUP BY g.game_name, pl.platform_name,
gp.release_year, pub.publisher_name
ORDER BY SUM(rs.num_sales) DESC;
```

Find the **average** number of sales of games released in **North America**

```
SELECT
AVG(rs.num_sales) AS
north_america_avg
FROM regional_sales rs
WHERE region_id = 1
```



HTML Application

We created a small HTML Website with our queries. Here are the answers to the 3 queries in our Demonstration, showcased in PostGre. There will be an index.html file in the final deliveries with all of our queries. There is also a demo available that will be shown during the presentation

How many **games** were released in **2016**

	count
1	344



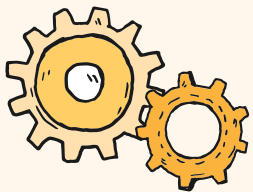
Group the games from **highest** selling

	game_name	platform_name	release_year	publisher_name	global_sales
	character varying (200)	character varying (200)	integer	character varying (200)	numeric
1	Wii Sports	Wii	2006	Nintendo	82.74
2	Super Mario Bros.	NES	1985	Nintendo	40.24
3	Mario Kart Wii	Wii	2008	Nintendo	35.83
4	Wii Sports Resort	Wii	2009	Nintendo	33.00
5	Pokemon Red/Pokemo...	GB	1996	Nintendo	31.38
6	Tetris	GB	1989	Nintendo	30.26
7	New Super Mario Bros.	DS	2006	Nintendo	30.01
8	Wii Play	Wii	2006	Nintendo	29.01
9	New Super Mario Bros....	Wii	2009	Nintendo	28.61
10	Duck Hunt	NES	1984	Nintendo	28.31
11	Nintendogs	DS	2005	Nintendo	24.75
12	Mario Kart DS	DS	2005	Nintendo	23.43
13	Pokemon Gold/Pokem...	GB	1999	Nintendo	23.09
14	Wii Fit	Wii	2007	Nintendo	22.72

Find the **average** number of sales of games released in **North America**

	north_america_avg
	numeric
1	0.2654666258420085





How many **games** released in **2016**?

```

 $\pi$  COUNT (release_year)
 $\gamma$  COUNT (release_year)
 $\sigma$  release_year = 2016 game_pl
  
```



Relational Algebra

Group the games from **highest** selling

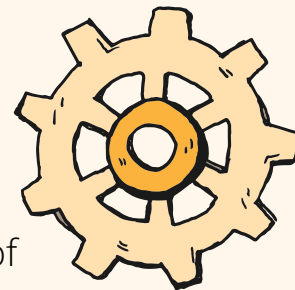
```

 $\tau$  SUM (num_sales)  $\downarrow$ 
 $\pi$  a.game_name, pl.platform_name, ap.release_year, pub.
publisher_name, SUM (num_sales)  $\rightarrow$  global_sales
 $\gamma$  game_name, platform_name, release_year, publisher_name,
SUM (num_sales)
( $\rho_{rs}$  regional_sales  $\bowtie$  rs.region_id = r.id
 $\rho_r$  region  $\bowtie$  rs.game_platform_id = ap.id
 $\rho_{ap}$  game_pl  $\bowtie$  ap.game_publisher_id = apub.id
 $\rho_{apub}$  game_publisher  $\bowtie$  apub.game_id = g.id
 $\rho_g$  games  $\bowtie$  ap.platform_id = pl.id
 $\rho_{pl}$  platform  $\bowtie$  apub.publisher_id = pub.id
 $\rho_{pub}$  publisher)
  
```

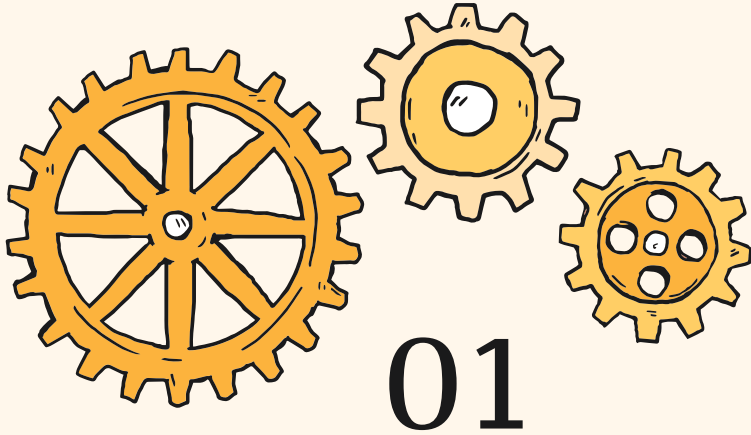
Find the **average** number of sales of games released in **North America**

```

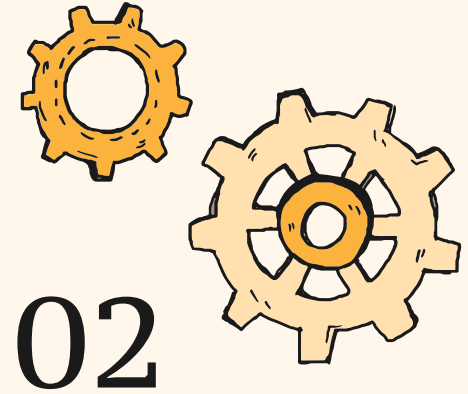
 $\pi$  AVG (num_sales)  $\rightarrow$  north_america_avg
 $\gamma$  AVG (num_sales)
 $\sigma$  region_id = 1
 $\rho_{rs}$  regional_sales
  
```



Conclusions



This project helped us to understand database and ER design a lot better. We learned databases are useful for many scenarios such as banks, pupils in a school, customer list etc



We feel our Structured Query Language skills (SQL), Python and HTML/CSS skills greatly improved during the designing of the database, report and presentation.